https://doi.org/10.3311/PPee.19755

Periodica Polytechnica Electrical Engineering and Computer Science, 66(2), pp. 105–115, 2022

Detection of Lung Cancer Stages on Computed Tomography Image Using Laplacian Filter and Marker Controlled Watershed Segmentation Technique

Tamanna Tajrin[1*], Mamun Ahmed[1], Sabina Zaman[1]

[1] Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Cumilla Cantonment, Cumilla, Bangladesh
[*] Corresponding author, e-mail: tamanna.tajrin@baiust.edu.bd

Received: 24 December 2021, Accepted: 05 April 2022, Published online: 27 April 2022

Abstract

Lung cancer is a form of malignant tumor distinguished by aggressive multiplication of abnormal cells in lung tissues. If we can assure the detection of lung cancer in the early stage, then we have a chance to increase the survival rate by five years as effective treatment is still available at this stage. Many researchers in the field of image processing sector have built various systems to detect cancer by using image processing techniques. Internationally TNM (Tumor, Nodule, Metastases respectively) method is followed by a physician and radiologist to describe the stage of lung cancer. Our proposed system uses image processing techniques to detect and classify the tumor according to the TNM staging method. First, a series of image processing techniques are performed in a Computed tomography (CT) image. Then, features are extracted to identify the region of interest (ROI). In our proposed system, the classification approach is different from the reviewed existing systems, and the detection rate is comparatively high.

Keywords

computed tomography (CT) image, image preprocessing, image segmentation, classification

1 Introduction

Lung cancer is the uncontrolled growth and division of cells. If the cells turn into a mass, then we refer to it as a tumor. Experts divide lung cancer into two types: benign and malignant. When the cancer term comes, it means we are pointing out malignant tumors that are growing aggressively. According to the survey information of the American Cancer Society, they have declared that most cancer deaths are caused by lung cancer. The known factors behind this disease are tobacco, excess body weight, hormones, and immune conditions. The survey conducted by the American Cancer Society also gave us more information, such as around 42% of newly diagnosed cases of cancer in the USA are preventable, whereas 19% of them are linked to smoking, and another 18% are caused by excess body weight, inactivity, and unbalanced nutrition. The treatment and survival rate of lung cancer is highly dependent on how quickly the small mass is identified [1]. In lung cancer, 85% are non-small cell, while the remaining 15% are small cell lung cancer. Further, we have studied six types of Non-small cell lung cancer. The types are as follows- occult, stage 0, stage 1, stage 2, stage 3, and stage 4. These stages are further subdivided by the tumor, nodes, and metastasis (TNM) system, which is mostly used by radiologists, pathologists, and doctors. Doctors combined their diagnosis results with pathological and radiological information to make sure whether cancer had spread from the lung field or not. In lung cancer, the death rate highly varies with stage. That is an exception to other cancer types. It is very important to stop or slow down the growth rate of the tumor to increase the patient's survival rate, and for that, we have to be more careful about the cancer stage with sub-staging information. Staging describes the spread of cancer during the diagnosis process. Proper staging is necessary to provide optimal treatment and prognosis. Nowadays, lung CT images are more effective for detecting and classifying cancer stages than chest X-rays [2]. It is anticipated that the cancer-related mortality rate will be 13% by 2030 in Bangladesh, and here lung cancer is leading the mortality rate upward day by day, which is a shocking fact for the country. The lifetime risk is higher for smokers than it is for non-smokers, but both are on the danger line. Several lung cancer treatments are available, such

Cite this article as: Tajrin, T., Ahmed, M., Zaman, S. "Detection of Lung Cancer Stages on Computed Tomography Image Using Laplacian Filter and Marker Controlled Watershed Segmentation Technique", Periodica Polytechnica Electrical Engineering and Computer Science, 66(2), pp. 105–115, 2022. https://doi.org/10.3311/PPee.19755

as chemotherapy, radiation therapy, biopsy, and surgery, but still, the five-year survival rate is only 14% [2]. The survival rate greatly depends on early detection and classification so that a system proves to be in demand for its efficacy in identifying cancer at its primary stage. Researchers from the USA, Europe, and Japan have worked a lot and are still working to develop new technology to detect lung cancer [1]. Hence, the purpose of this research is to develop a system that can both detect the presence of a tumor and define its stage by using various image enhancement and segmentation methods. To define the lung cancer stage, we have used an international standard that is followed worldwide by radiologists.

2 Literature review

The detection and classification system for lung cancer contains various steps, but the essential steps are image preprocessing, segmentation, feature extraction, and classification. Many algorithms have been developed to increase the detection rate of lung cancer detection systems. Some of those systems are discussed below.

Chaudhary and Singh [2], have proposed a system to identify lung cancer with the help of image processing techniques. They have used three main steps to develop their system. First, they have extended the quality of the CT image by using the Gabor filter. A marker controlled watershed segmentation technique was used to extract the partially expected portion of the enhanced image. Lastly, they have extracted some features like area, perimeter, and roundness to identify the stage of lung cancer. This system could not work well in the case of Gaussian noise; the Gabor parameter affects the output result in different ways for different CT images. The segmentation stage may fail to separate tumor and soft tissue in the case of real-time CT image.

Parveen and Kavitha [3], have proposed a system that analyzed the CT image and, most importantly, was an automated process that used various image processing techniques. Their methodology works in a sequential way as they have done image preprocessing by using a median filter. After that, they outlined the boundary of the lung and used a flood fill algorithm to extract the lung parenchyma. In the segmentation step, they have built an automated process where they first identified the threat pixel and then made an algorithm that can detect the tumor area from the threat pixel. At the last stage, the region-growing technique was used to detect the affected lung area. The authors limited their work to the preprocessing stage and didn't go for the feature extraction stage from a preprocessed image. The system cannot identify tumors that are close to the boundary of a lung field or connected with some parts of the lung, such as vessels and tissue, during the preprocessing stage.

Ignatious and Joseph [4] have introduced a computer-aided lung cancer detection system. They have reviewed three existing systems and identified the detection rate of the systems. They first reviewed a paper that used a neural fuzzy modelling technique where the thresholding method was used for segmenting CT images. A segmented image was used to extract features such as area, circularity, and brightness. Malignant CT images are separated from benign lung CT images based on extracted features. Then they reviewed another paper that proposed a system that used two steps for enhancement and segmentation of CT images. To remove noise from the image, they applied a median filter and a region-growing method used to achieve segmentation on CT images. The third paper introduced a lung cancer detection system with three essential steps. They used the Gabor filter to improve the quality of the lung CT image and the marker-controlled watershed segmentation technique to find the tumor region. Extracted features like area, perimeter, and eccentricity values are used to identify the cancer stage. This paper was a review-based paper, and we were not able to use it to detect stage 3 cancer images. If the tumor was adjacent to the lung field, then the preprocessing did not work well.

Abdillah et al. [5] developed a detection method that is totally dependent on image segmentation. They have used segmentation algorithms to identify lung cancer at its primary stage. For enhancement purposes, they used the Gabor filter, and after that, they used region-growing and marker control watershed segmentation techniques on the same image to find the mask. Then they used image binarization on that image. A pixel was assumed to be affected if its pixel value was larger than a specific threshold value. Otherwise, it was normal. In the end, this paper used the binarization method to determine the abnormality but didn't go to find the level of abnormality.

Wang and Wu [6] worked on a lung cancer detection system with image processing methods where they applied auto detection for finding tiny nodules and also developed a ridge detection algorithm. The diagnosis process was discussed and presented in this paper as a wavelet transform. Evaluation of the solidity of pulmonary nodules was done by using the decomposition technique, and to add new images, they used the subtraction algorithm. We have compared the algorithm with some of the traditional image segmentation algorithms. The authors actually focused on finding the tiny things within the lung field. They didn't try to extract features from the enhanced image.

Pratap and Chauhan [7] have invented a process to identify lung cancer at its primary stage. Their method was more about diagnosis at an early stage and a critical stage with intelligent computational techniques and various distortion removal algorithms. They converted the CT image to a gray scale image and then used some filters to remove noise from the image. Then they applied the threshold method and watershed technique to separate the desired portion. After that, they applied mathematical morphological operations on the resulted image to get the desired result. The authors limited their work to the preprocessing stage and didn't go for the feature extraction stage from a preprocessed image.

Kanazawa et al. [8], presented a method where they developed a system that can identify pulmonary nodules within helical CT images, and it was a computer-aided technique. Here they invented a computer-assisted diagnostic system that detects lung nodules at an early stage. They have separated blood vessels from the lung regions. The fuzzy model was used to make clusters, which helped the separation process. In some CT images, this method removed the fine details that can play an important role in denoting small tumor cells.

According to lung cancer researchers, helical CT is one of the most important CT modalities in the detection of pulmonary nodules. When we look at a CT image, we get a huge amount of data which has different features valued. So Armato et al. [9], developed a system that detected pulmonary nodules on CT images. This system involves two types of analysis: one was two-dimensional and the other one was three-dimensional. The rolling ball algorithm was used in the segmentation phase, which was done by a gray level threshold to get all types of nodules. Different threshold values in gray level have been used in the lung area to remove all the soft tissues and vessels except the lung regions. The noise reduction phase was not good enough to detect all types of pulmonary nodules.

Tariq et al. [10], developed a system to identify lung cancer. In that system, the threshold technique was used as a segmentation method. The post-processing phase starts only after the segmentation stage to identify the expected feature values in the CT image. Area, entropy, energy, eccentricity, standard deviation, and mean were the extracted features. Classification of cancer was done by using a neuro fuzzy classifier. If the tumor is adjacent to the lung field, then the preprocessing does not work well.

Song et al. [11] suggested a system that can identify the tumor automatically, and here the low level values have played an important role in detecting the abnormal lymph nodes and neighborhood features with Support Vector Machine (SVM). The SVM classifier was used with two-level conditions. One level was based on conditional random field (CRF) and the other level was based on relabeling the detected tumors by filtering the high uptake myocardium areas. The false detection rate was higher than what they have explained. We have not used the same dataset that they have used, so the rate may depend on the dataset.

Elizabeth et al. [12], proposed a system to identify lung cancer by using the greedy snake algorithm. The greedy snake algorithm is a segmentation algorithm that segments a CT image, and the segmented image was used to identify the ROI. The features of that region were extracted and used to detect cancer cells. There was no stage information for cancer to be detected by the system in the future.

Upon analyzing all the findings presented above, we have learned that a new and more efficient lung cancer detection system is necessary to identify cancer at its primary stage by applying the TNM staging method. Therefore, we have proposed a system that promises a higher detection rate than some of the popular existing systems. Our system can detect tumors that are close to the boundary of the lung field and work correctly even with Gaussian noise. Moreover, our system can predict the stage of cancer according to the T (tumor) size.

3 Methodology

The proposed system includes four steps: image preprocessing, feature extraction, tumor detection, and cancer stage classification, as in Fig. 1. The preprocessing stage has three phases: lung region extraction, region enhancement, and image segmentation. Each phase plays a significant role. In the first phase, we extracted the lung region from the original CT image. Then, in the second phase, we improved that extracted region, and in the last phase, we applied a segmentation method to the extracted lung region to find and separate the tumor.

3.1 Image section and data archive

We have taken axial-cut lung CT images in the proposed system, and the images are collected from Combined Military Hospital located in Cumilla cantonment, Bangladesh; Cumilla Medical College, Bangladesh; and

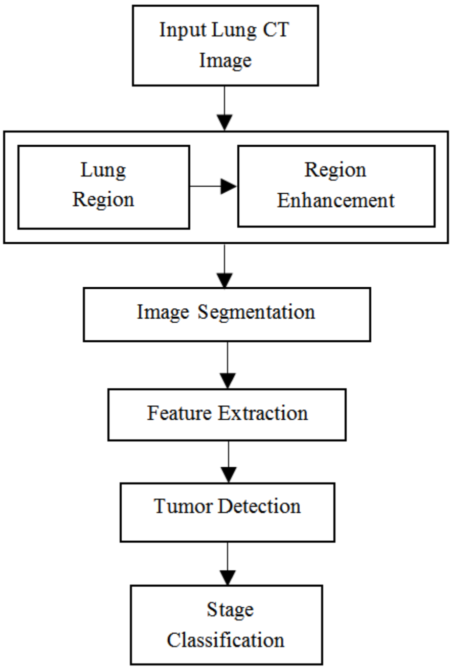

Fig. 1 Working flow of the proposed system.

The Cancer Imaging Archive (TCIA). Most of the detection systems use CT images in PNG or JPEG format, but we have used DICOM formatted lung CT images. To save the resultant image, we have converted the format from DICOM to JPEG. As a result, some of the intermediate phases may not be visually clear in the images, but the image from the last stage is enough to define the tumor.

3.2 Image pre-processing

Radiologists require lung cancer detection systems with high accuracy and performance rates so that they can detect cancer at an early stage with accurate staging information. In our proposed system, we first removed soft tissues from the lung CT image to increase the detection rate and then analyzed the edges and added them into the lung region. We have arranged sequential steps within the preprocessing stage to extract the lung area from the CT image and to enhance the edges within the lung region. Lung CT image enhancement, lung region extraction, and edge detection constitute a sequential task as described below with a flow chart in Fig. 2.

Noise removal is a very important phase in image processing, but for medical images, it is an unavoidable phase to get an accurate result. In our proposed system, we have applied the median filter to remove noise from the original lung CT image. Then get the sharp edges by applying the laplacian operator. The preprocessing steps of region extraction and edge detection are illustrated as follows:

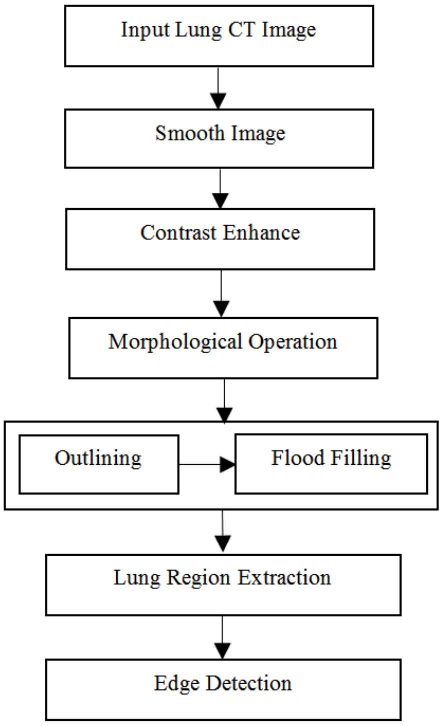

Fig. 2 Pre-processing phases of proposed system.

- Image Smoothing: The original lung CT image mostly contains salt and pepper noise. A noisy image can decrease the system's accuracy, so we have removed it by using a median filter. A median filter can smooth edges in an image so that a 3 × 3 kernel is better for the image, as a large-sized kernel may remove details from the image, which causes disturbance in further processing.
- Contrast Enhancement and Morphological Operation: The contrast level stretching technique is used in the proposed system because it helps a lot to get the desired image after performing morphological operations. Basically, erosion operations are done in the proposed system because if the tumor is adjacent to the lung border, then we have to separate it from the border, and this is done by erosion operations.
- Outlining and Flood Filling: Lung region outlining greatly depends on connected component property, which was measured by MATLAB's built-in function, and the extracted boundary was filled with white color, as we had selected it as our target color. To fill the lung region, we have applied a filling algorithm named "flood filling."
- Lung Region extraction: From the original image, we have extracted the lung region within the flood field area, from which we got only our ROI. The

extracted lung region came from the original image, so we needed an edge detector to enhance the inward edges.
- Edge detection: The Laplacian operator was used in the proposed system to extract edges from the image. Extracted edges are subtracted from the original image to enhance the edges within the lung region. In this case, it works like an unsharp mask. Preprocessing steps are shown in Fig. 3.

Watershed is an image segmentation method that is used to separate basins and assign a watershed to each separate region. A marker was applied to the watershed segmentation method and gave rise to a variant version named "Marker Controlled Watershed Segmentation Technique" because the marker was used to reduce over segmentation, which was the disadvantage of watershed segmentation. The total working process with a variant version of the watershed segmentation technique is stepped as:
- Step 1: Input the original lung CT image.
- Step 2: Search for a specific gradient magnitude.
- Step 3: Watershed transformation of gradient magnitude.
- Step 4: Opening Operations
- Step 5: Opening the image through the reconstruction method
- Step 6: Carry out the opening and closing operations.
- Step 7: Again, use the opening and closing through reconstruction method.

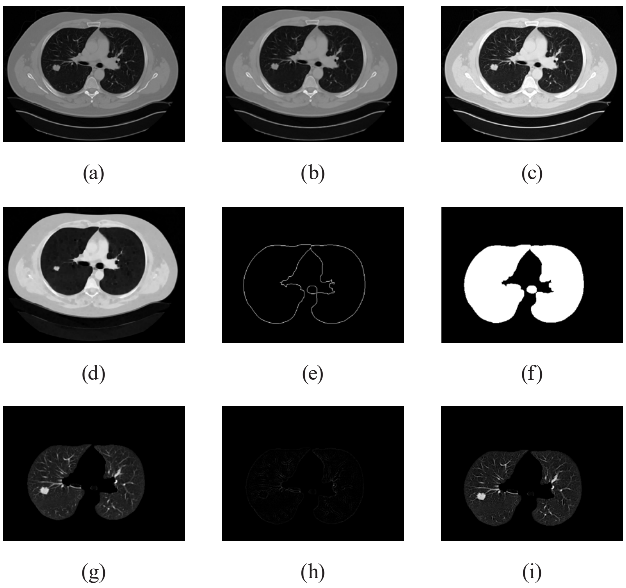

Fig. 3 Preprocessing steps (a) Original image; (b) Median filtered; (c) Contrast enhancement; (d) Erosion operation; (e) Outlining; (f) Flood filling; (g) Lung region extraction; (h) Edge detection; (i) Edge enhancement.

- Step 8: Determine the regional maxima of opening and closing using the reconstruction method.
- Step 9: Superimposed the resultant image from step8 on the original image.
- Step 10: Changed the regional maxima and superimposed them again on the original image.
- Step 11: Open and close the threshold using the reconstruction method.
- Step 12: Get the ridge lines.
- Step 13: Marker and object boundary placed together on the original image
- Step 14: Get the color matrix for the watershed label.
- Step 15: Colored labels are superimposed transparently on the original image.
- Step 16: Get the gray scale image.
- Step 17: Get the binary image.

Marker controlled watershed segmentation steps are applied in the developed system and the resultant images are shown in Fig. 4.

4 Feature extraction

After image preprocessing, we have extracted features of the suspected area. Based on the feature values, we can identify whether the area denotes a tumor or not. There are so many treatment procedures for lung cancer. These treatment procedures depend on the tumor location. We have applied an algorithm to detect and isolate various portions and shapes from a lung CT image. These features work as a basis for the classification stage. Feature extraction can influence the expected outputs to detect normality or abnormality, and if abnormal, then information about the stage of cancer is also extracted on the basis of features. We have considered four features in our proposed system to classify the cancer stages. Considered features are: area, perimeter, eccentricity, and diameter. Area and diameter are the most important features to identify the stage of lung cancer:
- Area: denotes the number of pixels with 1 and it's a scalar matrix [2]. We calculate the area value by adding all pixels that have 1.
- Perimeter: it's a scalar value. Here, perimeter describes the outline of an object. In a binary image, we can get the perimeter of an object by adding the interconnected pixels [2].
- Eccentricity: it is also referred to as roundness. It is also known as the roundness, circularity, and irregularity matrix. This matric value is set to 1 only for

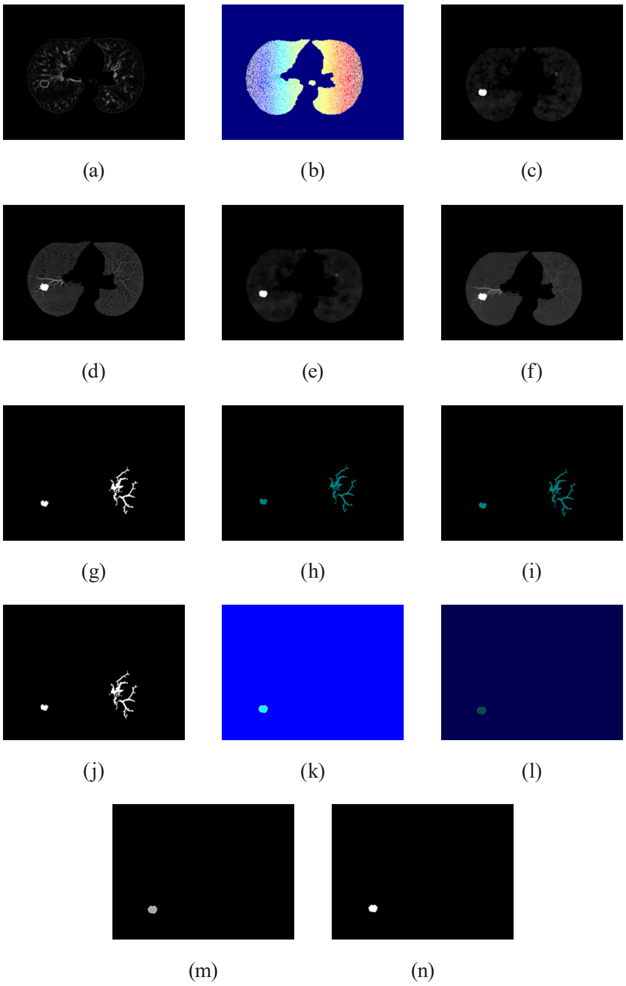

Fig. 4 Resultant images (a) Gradient magnitude of image; (b) Watershed transform of gradient magnitude; (c) Opening; (d) Opening the image through reconstruction method; (e) Opening and Closing operation on the image of step 5; (f) Opening and Closing through reconstruction method; (g) Regional maxima of Opening and Closing; (h) Superimposed the resultant image of step8 on original image; (i) Changed regional maxima and again superimposed on original image; (j) Marker and object boundary placed together on original image; (k) Color matrix for the watershed label; (l) Color labels are placed on original image; (m) Grey scale image; (n) Binary image.

circular shapes, and it is < 1 for any other shapes [2]. In our proposed system, we consider the object's more circularity.

- Diameter: Lung tumors are typically circular in shape, and radiologists determine the stage of lung cancer based on their diameter. To find the diameter, we have considered the distance between the two furthest points of the tumor. In our proposed system, we give greater importance to diameter as it helps in the auto staging part of our system.

5 Experimental result

The proposed system was tested with axial-cut lung CT images, which were collected from different sources. The dataset includes 40 patients and 6000 slices of 512 × 512 pixels in size. Physicians examined all the slices and found 37 patients with lung tumors. The proposed system identified 36 patients with abnormal lungs and predicted the cancer stage, which matched with the physicians' report:

- T1: a tumor that is less than or equal to 3 cm in diameter, is surrounded by lung/visceral pleura, and does not involve the main bronchus.
- T2: a tumor measuring 3 to 5 cm in diameter or involvement of the main bronchus without a carina.
- T3: a tumor larger than 5 to 7 cm in diameter or a tumor of any size involving the chest wall, pericardium, phrenic nerve, or satellite nodules in the same lobe.
- T4: We can assume that the tumor has spread to other parts of the body at this point.

At the end of the preprocessing, we identify the area of each object and select only the objects that fulfill our predefined eccentricity value. By this, we got the tumor within the lung field. Then we measured the perimeter of the tumor and calculated the diameter. Finally, we took the tumor's diameter into account and determined the lung cancer stage. Extracted features of 20 patients are displayed in Table 1.

From Table 1, we can say that the proposed system worked well to identify stages 1, 2, and 3a. After getting this expected result from our system, we discussed it with two radiologists. Radiologists checked the CT images that we have used in our system and gave their opinions on those images. We have compared their opinion with the result that we got from the proposed system. In Table 2, we have shown 20 patients' lung cancer staging information that was obtained from physicians and identified by the proposed system.

We will show you in Section 6 that our detection rate is higher than other methods. We will also discuss the present systems' flaws and how we addressed them.

6 Discussion

In our research, we have compared various preprocessing methods to find an effective method that works well with the segmentation technique. We have worked with

Table 1 Extracted features and predicted lung cancer stages.

Area	Eccentricity	Diameter	Perimeter	Stage
1081	0.38661	37.099	132.35	1
1257	0.67322	40.006	207.57	2
1429	0.88657	42.655	190.59	2
138	0.65881	13.255	46.735	1
1808	0.90313	151.73	767.58	4
1658	0.60768	45.946	154.66	1
267	0.87578	18.438	77.426	1
418	0.54017	23.07	73.654	1
1010	0.80024	35.86	226.88	1
497	0.71477	25.156	87.094	1
1576	0.74324	44.795	184.02	2
1335	0.76448	130.4	618.9	4
1512	0.36916	138.75	581.76	4
366	0.86592	21.587	77.616	1
1510	0.68675	43.847	240.04	2
3316	0.59722	64.977	229.42	3
222	0.6156	16.812	54.021	1
82	0.6363	10.218	66.542	1
4536	0.61684	75.996	344.06	2
2408	0.75682	55.371	465.68	2
182	0.59456	15.223	46.991	1
1540	0.66128	44.281	393.88	2
816	0.85863	32.233	159.01	1
968	0.78949	35.107	156.95	1
366	0.78485	21.587	72.455	1
853	0.86392	32.956	188.21	1
278	0.81017	18.814	65.789	1
651	0.90085	28.79	310	1
74	0.48211	9.7067	61.844	1
3357	0.74597	65.378	220.49	2
805	0.94709	32.015	208.68	1
218	0.92336	16.66	79.543	1
472	0.86205	24.515	155.65	1
300	0.97264	19.544	117.86	1
3057	0.95597	60.254	200.35	2
574	0.80423	27.034	127.51	1
4033	0.58943	72.593	323.76	2
2108	0.78253	52.392	487.93	2
1396	0.86559	42.16	279.34	2
455	0.81586	24.069	98.067	1
583	0.84889	27.245	136.04	1

different filters like- median, Gaussian, Gabor, auto level, histogram, morphological operator, and Laplacian operator. The median filter makes the edges of an image smooth or blur, so we need to apply some techniques that enhance the edges within the CT image. We can enhance images either in the spatial domain or in the frequency domain [13]. In that case, the laplacian filter becomes a proper operator as it can enhance edges in the CT image. The existing lung cancer identification systems are highly dependent on the segmentation technique, so the identification rate may fluctuate with different segmentation techniques. We have also compared several segmentation algorithms, such as thresholding, region growing, adaptive thresholding, watershed segmentation, and marker controlled watershed segmentation techniques. A marker is used in the watershed segmentation method to get an improved version of the watershed method, named the marker-controlled watershed segmentation method. By using the marker, we can remove an unnecessary segmentation problem from the original watershed segmentation method. Different lung cancer detection systems use various combinations of preprocessing methods and segmentation algorithms in Table 3.

Table 2 Lung cancers staging by physicians and proposed system.

Patient number	Clinical staging	Staging by proposed system
Patient 1	1A	1
Patient 2	2B	2
Patient 3	2A	2
Patient 4	1B	1
Patient 5	1A	1
Patient 6	1A	1
Patient 7	1B	1
Patient 8	1C	1
Patient 9	1B	1
Patient 10	2B	2
Patient 11	3B	3
Patient 12	3A	3
Patient 13	1A	1
Patient 14	2B	2
Patient 15	3A	3
Patient 16	1B	1
Patient 17	1A	1
Patient 18	2A	2
Patient 19	2A	2
Patient 20	1A	1

A neural fuzzy model [1] can identify pulmonary nodules within the lung CT image. Lin and Yan [1] applied the thresholding, morphological, and labeling operations to get the lung area from the original image. Then identify the region of interest and extract three features such as area, brightness, and circularity from that region.

Table 3 Lung cancer detection systems with various algorithms used in different stage.

System	Enhancement stage	Segmentation stage	Feature extraction stage
Neural fuzzy model [1]	–	Thresholding	Area, circularity, brightness
Region growing method [3]	Median filter	Region growing	–
Lung cancer detection on CT image [2]	Convolution filters with Gaussian pulse, Gabor filter	Marker controlled watershed segmentation technique	Area, eccentricity, perimeter

Diagnose rules are used to identify the nodules where the rules are obtained from the neural fuzzy model as they have mentioned.

The automatic region growing method [3] can identify pulmonary nodules and is applied in a system where a median filter is used in the pre-processing stage, and threat point identification is applied with the region growing method to separate a defected region. This system does not extract any features to identify the cancer stage.

A system [2] used image processing steps to identify the lung tumor from the lung CT image, where they used some image enhancing filters like the Convolution and Gabor filters. After enhancing the image, they used an improved version of the watershed segmentation method to get expected features like perimeter area and eccentricity of the tumor region.

We have implemented some existing systems and found some limitations in the case of real-time CT images.

In a threshold-based lung cancer detection system [1], we have found two major limitations: It can detect only those nodules that have a diameter of between 1 cm and 5 cm within the lung field, and this system may fail to detect nodules that are close to the boundary of the lung field during the segmentation stage. These limitations are pointed out through images and displaying in Fig. 5.

In the region-growing method-based lung cancer detection system [3], we have found a shortcoming and it is as follows. The system cannot identify tumors that are close to the boundary of a lung field or connected with some parts of the lung, such as vessels and tissue, during the preprocessing stage.

This limitation is pointed out after implementing CT images in real-time, as shown in Fig. 6.

The marker control watershed method-based lung cancer detection system [2] has limitations as follows. It cannot work well in the case of Gaussian noise, the Gabor parameter effects the output result in different ways for different CT images, and the segmentation stage may fail to separate tumor and soft tissue in the case of real-time CT image. These limitations are pointed out after implementing CT images in real time and are displayed below in Fig. 7.

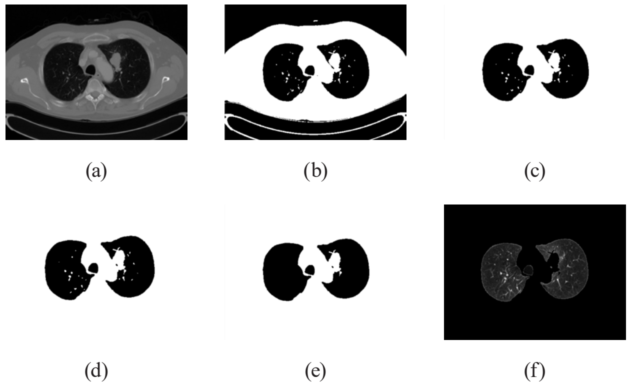

Fig. 5 Pointed out limitations through images (a) The input image; (b) Thresholded; (c) Extracted lung area; (d) Morphological operation (closing); (e) Filling; (f) Separated lung area.

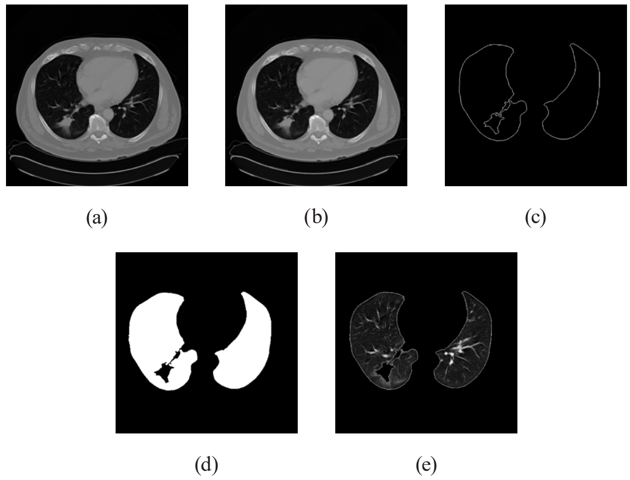

Fig. 6 Pointed out limitation after implementing CT images in real-time (a) Input CT image; (b) Applied filter; (c) Draw the border; (d) Filling process; (e) Separated lung area.

Some of the above-implemented systems did not work on the feature extraction step, which is an important step in identifying cancer at its primary stage.

We overcame the shortcomings of the existing system and added some new techniques to our proposed system, which we mentioned earlier in this paper. The major shortcomings that we have removed from the existing systems are as follows:

1. Our system can detect tumors that are close to the boundary of the lung field.
2. It can work well in the case of Gaussian noise.

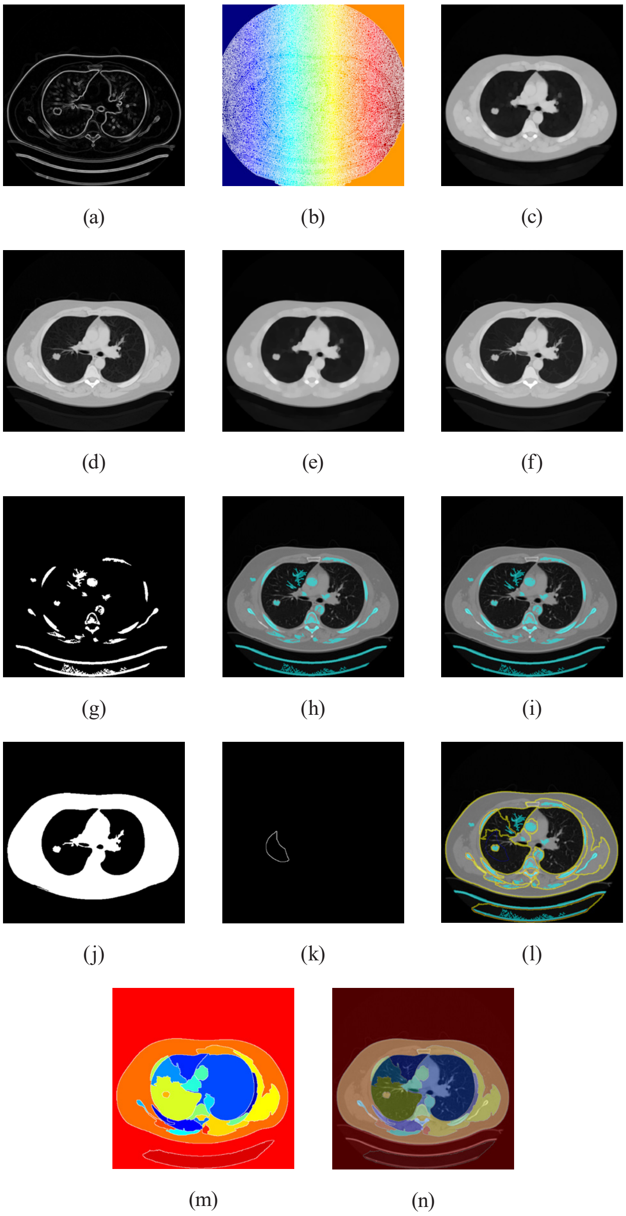

Fig. 7 Pointed out limitations after implementing CT images in real time (a) Gradient image; (b) Watershed transformation; (c) Opening; (d) Reconstruction for opening; (e) Morphological Operation (opening-closing); (f) Morphological operation by reconstruction; (g) Search regional maxima; (h) Regional maxima placed on the image; (i) Change the regional maxima; (j) Thresholding and reconstruction; (k) Show the ridge line; (l) Markers and boundaries; (m) Label matrix; (n) Label matrix on inputted image.

3. The segmentation stage has succeeded in separating tumor and soft tissue in the case of real-time CT imaging.

Our preprocessing steps helped us to achieve the major contributions that we have pointed out in the previous paragraph. In addition, we have worked on features to detect the stage of lung cancer. The way we chose the tools helped us to increase the detection rate, which we have mentioned in Section 3 and also in Section 6. We have proposed a new system where we first separated lung fields from the original image, then enhanced inward edges and separated tumors from vessels and soft tissues by applying the marker control watershed method. This proposed system can automatically define the lung cancer stages on the basis of conditions that work on extracted features. The detection rate of the system is determined by using the following formula:

$$\text{Detection Rate} = (a-b)/a \times 100\%, \qquad (1)$$

where:

- a: no. of total patients,
- b: no. of incorrect output.

We have calculated the detection rates of some existing systems and our proposed systems by using real time images; consulted with radiologist and based on the radiologist the detection rates are given in Table 4.

In our testing process, the neural fuzzy model [1] failed to identify 9 lung cancer affected CT images, and the detection rate was 77.5%, as shown in Table 1 above. With 82.5% detection rates, the region-growing method-based system [3] failed to identify 7 lung cancer affected CT images and, with 85% detection rates, the lung cancer detection on CT images using an image processing system [2] failed to detect 6 lung cancers affected CT images.

From Table 4 and Fig. 8, it is clear that the proposed system provides an optimum solution to identify lung tumors in their primary phase with 90% detection rates.

The detection rate varies with preprocessing techniques and also with classification methods. We got some research work, where researchers applied various classification techniques with different preprocessing methods and got multiple results [14, 15]. The preprocessing method was unchanged, but they changed the classifier and got better results. Sometimes, classification techniques play a very important role in identifying the cancer stage.

Table 4 Detection rates of different systems.

System	Detection rate
Neural fuzzy model [1]	77.5%
Region growing method [3]	82.5%
Lung cancer detection on CT image using image processing [2]	85%
Proposed system	90%

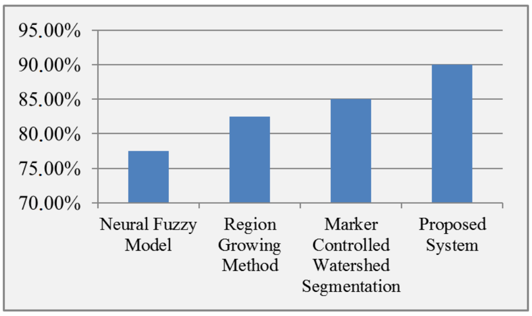

Fig. 8 Graphical representation of detection rate of various systems.

7 Conclusion

The proposed system is an effective lung cancer detection system that can identify tumors and define the stage of cancer with higher accuracy at its early stage. The resulting lung CT image is preprocessed by using a series of methods, including median filtering, contrast enhancing, morphological operation, outlining, flood filling, and laplacian operator. At the end of the preprocessing stage, we obtained the lung field from the CT image and applied a different version of the watershed segmentation technique, named the marker control watershed segmentation technique, to identify tumors within a region of interest. We have extracted four features they are area, perimeter, eccentricity, and diameter of the tumor. This combination leads our system to define the cancer stage. The detection rate of the proposed system is 90% and can help radiologists detect and define the lung cancer stage at its primary stage.

In the near future, we will implement different types of classifiers to increase the detection rate of our proposed system.

Acknowledgement

At the most, we are thankful to the Radiology Department of Cumilla Medical College, Bangladesh, for providing us with the required information and also for sharing their knowledge and insights. We would like to express our thanks to the Combined Military Hospital located in Cumilla cantonment, Bangladesh for giving us valuable data that we used to develop our system.

References

[1] Lin, D. T., Yan, C. R. "Lung nodules identification rules extraction with neural fuzzy network", In: Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02., Singapore, Singapore, 2002, pp. 2049–2053. https://doi.org/10.1109/ICONIP.2002.1199035

[2] Chaudhary, A., Singh, S. S. "Lung Cancer Detection on CT Images by Using Image Processing", In: 2012 International Conference on Computing Sciences, Phagwara, India, 2012, pp. 142–146. https://doi.org/10.1109/ICCs.2012.43

[3] Parveen, S. S., Kavitha, C. "Detection of lung cancer nodules using automatic region growing method", In: 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 2013, pp. 1–6. https://doi.org/10.1109/ICCCNT.2013.6726669

[4] Ignatious, S., Joseph, R. "Computer aided lung cancer detection system", In: 2015 Global Conference on Communication Technologies (GCCT), Thuckalay, India, 2015, pp. 555–558. https://doi.org/10.1109/GCCT.2015.7342723

[5] Abdillah, B., Bustamam, A., Sarwinda, D. "Image processing based detection of lung cancer on CT scan images", Journal of Physics: Conference Series, 893, Article number: 012063, 2017. https://doi.org/10.1088/1742-6596/893/1/012063

[6] Wang, W., Wu, S. "A Study on Lung Cancer Detection by Image Processing", In: 2006 International Conference on Communications, Circuits and Systems, Guilin, China, 2006, pp. 371–374. https://doi.org/10.1109/ICCCAS.2006.284656

[7] Pratap, G. P., Chauhan, R. P. "Detection of Lung cancer cells using image processing techniques", In: 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 2016, pp. 1–6. https://doi.org/10.1109/ICPEICES.2016.7853347

[8] Kanazawa, K., Kawata, Y., Niki, N., Satoh, H., Ohmatsu, H., Kakinuma, R., Kaneko, M., Eguchi, K., Moriyama, N. "Computer-aided diagnosis for pulmonary nodules based on helical CT images", In: 1997 IEEE Nuclear Science Symposium Conference Record, Albuquerque, NM, USA, 1997, pp. 1635–1639. https://doi.org/10.1109/NSSMIC.1997.670631

[9] Armato, S. G., Giger, M. L., Moran, C. J., Blackburn, J. T., Doi, K., MacMahon, H. "Computerized Detection of Pulmonary Nodules on CT Scans", RadioGraphics, 19(5), pp. 1303–1311, 1999. https://doi.org/10.1148/radiographics.19.5.g99se181303

[10] Tariq, A., Akram, M. U. Javed, M. Y. "Lung nodule detection in CT images using neuro fuzzy classifier", In: 2013 Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI), Singapore, Singapore, 2013, pp. 49–53. https://doi.org/10.1109/CIMI.2013.6583857

[11] Song, Y., Cai, W., Kim, J., Feng, D. D. "A Multistage Discriminative Model for Tumor and Lymph Node Detection in Thoracic Images", IEEE Transactions on Medical Imaging, 31(5), pp. 1061–1075, 2012. https://doi.org/10.1109/TMI.2012.2185057

[12] Elizabeth, D. S., Nehemiah, H. K., Raj, C. S. R., Kannan, A. "Computer-aided diagnosis of lung cancer based on analysis of the significant slice of chest computed tomography image", IET Image Processing, 6(6), pp. 697–705, 2012. https://doi.org/10.1049/iet-ipr.2010.0521

[13] Gonzalez, R. C., Woods, R. E. "Digital Image Processing", Pearson Prentice Hall, Upper Saddle River, NJ, USA, 2008.

[14] Senthil Kumar, K., Venkatalakshmi, K., Karthikeyan, K. "Lung Cancer Detection Using Image Segmentation by means of Various Evolutionary Algorithms", Computational and Mathematical Methods in Medicine, 2019, Article ID: 4909846, 2019. https://doi.org/10.1155/2019/4909846

[15] Shakeel, P. M., Burhanuddin, M. A., Desa, M. I. "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier", Neural Computing and Applications, 2020. https://doi.org/10.1007/s00521-020-04842-6